

## Article

# Can Genetic Testing Identify Talent for Sport?

Pickerin, Craig, Kiely, John, Grgic, Jozo, Lucia, Alejandro and Del Coso, Juan

Available at <http://clock.uclan.ac.uk/30910/>

*Pickerin, Craig, Kiely, John ORCID: 0000-0001-9817-0224, Grgic, Jozo, Lucia, Alejandro and Del Coso, Juan (2019) Can Genetic Testing Identify Talent for Sport? Genes, 10 (12). e972. ISSN 2073-4425*

It is advisable to refer to the publisher's version if you intend to cite from the work.

<http://dx.doi.org/10.3390/genes10120972>

For more information about UCLan's research in this area go to <http://www.uclan.ac.uk/researchgroups/> and search for <name of research Group>.

For information about Research generally at UCLan please go to <http://www.uclan.ac.uk/research/>

All outputs in CLoK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the <http://clock.uclan.ac.uk/policies/>

# Can Genetic Testing Identify Talent for Sport?

Craig Pickering <sup>1</sup>, John Kiely <sup>1</sup>, Jozo Grgic <sup>2</sup>, Alejandro Lucia <sup>3,4</sup> and Juan Del Coso <sup>5,\*</sup>

<sup>1</sup> Institute of Coaching and Performance, School of Sport and Wellbeing, University of Central Lancashire, Preston PR1 2HE, UK; craigpickering1014@hotmail.com (C.P.); jkiely@uclan.ac.uk (J.K.)

<sup>2</sup> Institute for Health and Sport (IHES), Victoria University, Melbourne, VIC 3011, Australia; jozo.grgic@live.vu.edu.au

<sup>3</sup> Faculty of Sport Sciences, Universidad Europea de Madrid, Villaviciosa de Odón, 28670 Madrid, Spain; alejandro.lucia@universidadeuropea.es

<sup>4</sup> Research Institute i+12, and Centro de Investigación Biomédica en Red de Fragilidad y Envejecimiento Saludable, 28041 Madrid, Spain

<sup>5</sup> Centre for Sport Studies, Rey Juan Carlos University, 28943 Fuenlabrada, Spain

\* Correspondence: juan.delcoso@urjc.es

Received: 4 November 2019; Accepted: 23 November 2019; Published: 26 November 2019

**Abstract:** Elite athlete status is a partially heritable trait, as are many of the underpinning physiological, anthropometrical, and psychological traits that contribute to elite performance. In recent years, our understanding of the specific genetic variants that contribute to these traits has grown, such that there is considerable interest in attempting to utilise genetic information as a tool to predict future elite athlete status. In this review, we explore the extent of the genetic influence on the making of a sporting champion and we describe issues which, at present, hamper the utility of genetic testing in identifying future elite performers. We build on this by exploring what further knowledge is required to enhance this process, including a reflection on the potential learnings from the use of genetics as a disease prediction tool. Finally, we discuss ways in which genetic information may hold utility within elite sport in the future, including guiding nutritional and training recommendations, and assisting in the prevention of injury. Whilst genetic testing has the potential to assist in the identification of future talented performers, genetic tests should be combined with other tools to obtain an accurate identification of those athletes predisposed to succeed in sport. The use of total genotype scores, composed of a high number of performance-enhancing polymorphisms, will likely be one of the best strategies in the utilisation of genetic information to identify talent in sport.

**Keywords:** genetics; single nucleotide polymorphism; athletic performance; elite athlete; success

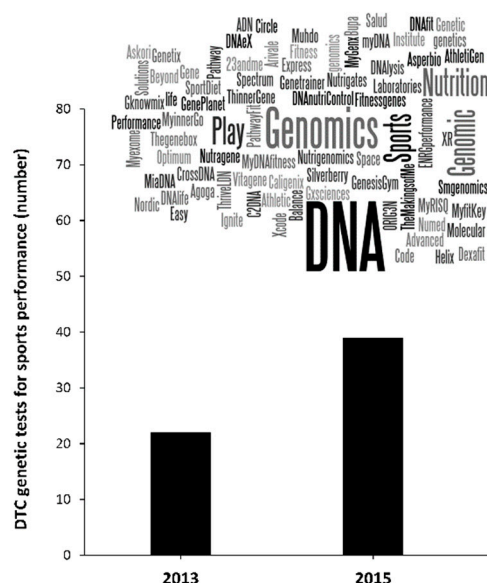
## 1. Introduction

It is widely recognised that some individuals are naturally talented in specific physical traits that are, in turn, related to sports performance [1]. In some cases, talented individuals come from talented families, suggesting that genetics is partially responsible for the physical, physiological or anthropometric characteristics that are needed for attaining athletic success [2]. Indeed, elite athlete status is partially heritable, with twin studies suggesting that 30–80% of the variance in this trait is explained by heritable factors [3]. Recent advances in genetic technology have allowed for greater exploration of the genetic underpinnings of elite performance, leading to the identification of single nucleotide polymorphisms (SNPs) and other genetic variants with the potential to influence sports performance, either directly or indirectly. For example, a SNP in *ACTN3*, R577X (rs1815739), has been shown to modify the attainment of elite speed–power athlete status [4,5]. Here, a common C-to-T base substitution results in the transformation of an arginine base (R) to a premature stop codon (X). X allele homozygotes are deficient in the protein encoded for by *ACTN3*,  $\alpha$ -actinin-3, which is

expressed exclusively in fast twitch muscle fibres. As a result, XX genotypes tend to have lower proportions of fast-twitch muscle fibres [6] and, because fast-twitch muscle fibres are an important component of speed–power performance, these genotypes might be underrepresented in elite speed–power cohorts [7]. The first study on this topic was conducted by Yang and colleagues [7], who reported that the X allele was significantly less frequent in elite male and female sprint athletes than in non-athletic controls and elite endurance athletes. Whilst in Caucasian populations the frequency of the XX genotype is ~20%, in Yang and colleague’s [7] cohort of Caucasian power Olympians, the XX genotype was entirely absent. The finding of significantly lower frequencies of the X allele and XX genotypes in elite speed–power athletes has been replicated [8,9], although equivocal findings have also been reported elsewhere [10]. Consequently, despite generally being estimated to explain around 1–3% of the variance in speed–power phenotype [2,11,12], *ACTN3* has subsequently been labelled a “gene for speed” [13–15]. *ACTN3*, however, is not the only gene associated with elite athlete status, with a recent review reporting that at least 155 genetic markers are potentially linked to elite athlete status, at least in some cohorts or ethnicities [16].

Whilst many of the currently established SNPs associated with elite athlete status are associated with physiological traits such as aerobic capacity, muscle strength or speed, there is the potential that other SNPs may exert a less direct—but no less important—effect on the attainment of elite sports performance. For example, both height [17] and body mass index are highly heritable [18], and both likely contribute to the achievement of elite athlete status on a sport-by-sport basis. Furthermore, psychological traits are also genetically influenced, with a number of SNPs associated with anxiety [19]. Interestingly, it has been found that variants in genes coding for proteins potentially modulating activity of brain emotion centers, particularly in genes encoding elements of serotonergic, catecholaminergic and hypothalamic–pituitary–adrenal systems might also influence the predisposition to elite sports performance [20]. As such, the genetic influence on sporting performance is broad, multi-factorial, and pervasive.

Given the wide-ranging and potential influence of genetic variations on the attainment of elite athlete status, there is considerable interest in collecting genetic information to identify athletes with the potential to achieve elite status. Indeed, the number of companies that offer direct-to-consumer (DTC) genetic testing to predict sports performance/injury risk associated with sport has rocketed in recent years (Figure 1). Moreover, contemporary reports detail the use of genetic testing within the talent identification process in a number of countries and for a wide-range of sport enthusiasts [21]. One of the concerns with the use of such DTC genetic testing is that the information provided by the companies is based on a very limited genetic analysis, while the misuse of the research evidence to endorse the utility of this type of testing for early talent selection is common. The pace of growth in the number of companies that offer DTC genetic testing occurs partially due to the absence of worldwide agreements for a common regulation on this type of genetic testing. For these reasons, the general consensus amongst researchers is that these tests should not be used for talent identification [21], and the use of such tests in this way is ethically troubling [22]. Here, we discuss why current genetic tests cannot predict future sporting success, and explore the necessary advancements required to enable the utilisation of genetic information to accurately identify future talented performers.



**Figure 1.** Number of companies that offer direct-to-consumer (DTC) genetic testing marketed as being related to sport performance, exercise performance and sports injury risk. Data in 2019 has been obtained by using the search terms “genetic”, “test”, “exercise” and “sport” in two popular search engines (Google and Bing) replicating the procedures followed by Williams et al., in 2013 [23] and Webbort et al., in 2015 [21].

## 2. Why We Cannot Currently Use Genetic Information for Talent ID

As previously mentioned, over 155 genetic markers have been tentatively linked with elite athlete status [16], although the association of most of these genetic variants to sports performance has a weak scientific background. The various suggestive genetic variants are typically—but not always—divergent, such that they predispose towards an increased chance of success in either power–strength or endurance sports/events, but not both. These divergent effects demonstrate that there is not a singular genetic profile that confers sporting success, but that the required genetic profiles are likely specific for characteristics of each sport. Whilst some of these markers, such as the aforementioned *ACTN3* rs1815739 [5,7], and to a lesser extent angiotensin-converting enzyme (*ACE*) insertion (I)/deletion (D) polymorphism (rs1799752) [5,24], and peroxisome proliferator-activated receptor gamma coactivator 1-alpha (*PPARGC1A* or simply *PGC1-1α*) rs8192678 [25,26] have been replicated in some studies; others, such as transcription factor A, mitochondrial (*TFAM*) rs1937, have yet to be satisfactorily replicated [27]. Of the 155 genetic variants identified as linked to elite athlete status [16], 31 have been associated with elite athlete status in at least two studies, with 12 replicated in at least three [16]. Accordingly, at present, we know only a few of the genetic markers that likely associate with elite athlete status, making predictions of future sporting prowess based on such information not only difficult but also likely inaccurate. In addition, testing the utility and efficacy of genetic testing in the talent identification process should be conducted to ascertain whether the information provided by genetic testing—and not obtained through other traditional non-genetic tests such as physical testing—is of relevance to increase the specificity of overall talent selection [28].

Another issue is that the currently available markers appear to offer poor specificity and sensitivity as talent identification tools. Returning to *ACTN3* rs1815739, whilst the R allele is associated with elite athlete status in speed–power sports events [5], with the XX genotype potentially playing a negative role in performance in such types of events [7], it remains unclear how discriminatory this information might be. In Caucasians, for example, ~80% of individuals possess an R allele [7]. In some black African populations, this percentage can be as high as 99% [29]. In a study of elite US and Jamaican sprinters—the populations providing the fastest eight 100 m runners of all time—there was no significant difference in *ACTN3* genotype frequencies between these athletes and their non-athletic peers, with 97% of non-athletes possessing at least one R allele [10].

Therefore, whilst the R allele may be required for elite sprint performance, as the vast majority of the world's population possess it, knowledge of an individual's *ACTN3* genotype is not particularly useful from a talent identification perspective. Furthermore, there are exceptions to the belief that an R allele is required for elite speed–power performance. In their study of elite sprinters, Papadimitriou and colleagues [12] reported that one male and one female 100 m sprinter, both of whom achieved the Olympic qualifying standard, did not possess an R allele. Additionally, Lucia and colleagues [30] reported the case of a long-jumper with a personal best of 8.26 cm, just 5 cm off the gold medal winning jump at the 2012 Olympic Games, who had the XX genotype. Such findings demonstrate that the lack of an *ACTN3* R allele does not preclude attainment of elite status in speed–power events. Additionally, performance-enhancing polymorphisms may still have a low prevalence in elite athlete populations. For example, a SNP (rs7181866) in nuclear respiratory factor 2 (*NRF2*) gene has been associated with elite athlete status, with a significantly higher proportion of the AG genotype compared to the AA genotype found in elite endurance athletes when compared to controls [9]. However, only 12–14% of these elite athletes possessed the “ideal” AG genotype, illustrating that the vast majority of elite athletes did not have this specific performance-enhancing polymorphism, again potentially limiting its use as a discriminating screen.

Such findings demonstrate the difficulties of a single gene approach for talent identification, highlighting that such an approach is no longer viable. In light of these findings, researchers have turned to Total Genotype Scores (TGS), whereby a number of elite athlete-associated SNPs are combined into a single polygenic score. Ruiz and colleagues [31,32] utilised this approach involving elite endurance and power athletes. For their study focusing on endurance athletes, they combined seven polymorphisms into a total score, finding that the mean score was higher in the athlete group than in the control (non-athletic) group [31]. This finding was replicated using a TGS comprised of six SNPs for power athlete status, with elite power athletes having a higher score than both endurance athletes and non-athletic controls [32]. Possession of the “perfect” polygenic profile (i.e., the elite athlete genotype of all SNPs) was rare, occurring in only 9.4% of the power athletes, and in no endurance athletes. In addition, there was considerable overlap between groups, such that a number of controls had better TGS than some elite power and endurance athletes [32]. Furthermore, whilst a TGS may help in discriminating between athlete and non-athlete, Santiago and colleagues [33] demonstrated that, in a group of rowers, it did not distinguish between different levels of performance (i.e., World vs National medallists). Earlier work [34,35] demonstrates that there is considerable similarity in polygenic scores within humans—athletes and non-athletes alike—when a low number of markers (in this case, 22 and 23) are used, such that, again, this approach would likely have limited real world specificity and sensitivity. Lastly, the configuration of TGS to differentiate between populations of athletes usually confers the same score to each SNP while the individual influence of each gene to the status of elite athlete is likely uneven, with some genetic variants exerting a greater influence than others. In order to improve the insights provided, a far greater number of performance-enhancing polymorphisms are likely required, whilst the relative weighing of the influence of each SNP may add further value in the development of increasingly accurate TGS. As such, further work should seek to identify additional SNPs associated with elite athlete status across a variety of sporting phenotypes, and replicate existing SNPs such that there is a pool of SNPs associated with elite athlete status that researchers have confidence in. The relative contribution of each SNP towards elite athlete status could then be determined via assessing the predictive ability of differently weighted algorithms, which can then be tested in hold-out cohorts for validation.

As a summary of the above discussion, it seems clear that the provision of elite athlete status is a highly complex, polygenic trait, and that we know very few of the genetic variations that contribute to such a trait. As a result, it appears a fundamental requirement that, if genetic testing is to be utilised for talent identification purposes, a far greater number of polymorphisms associated with elite athlete status needs to be uncovered and then combined into a complex TGS model.

### 3. What Further Knowledge Do We Require to Potentially Use Genetic Information for Talent Identification?

### 3.1. Genome-Wide Association Studies

The evolution of Genome-Wide Association Study (GWAS) methodology potentially offers an opportunity to expand the number of genetic variants currently associated with elite athlete status, or to other traits associated with sporting success, such as the likelihood of suffering performance-limiting injuries [36]. Whilst the majority of the SNPs currently associated with elite athlete status were elucidated via gene-association studies or candidate gene analysis—where a limited number of SNPs are hypothesised to have an effect on some particular trait associated with exercise performance, and that hypothesis is tested—GWAS are hypothesis-free. In a GWAS, a large number of SNPs are analysed for association with a specific trait. Given that there is no hypothesis to be tested, it provides a robust method to detect novel associations. However, due to the very low p-values required to reach genome-wide significance ( $p < 5 \times 10^{-8}$ ), and the (often) very low effect sizes of any SNP, GWAS studies require very large subject numbers and might fail to identify some variations with a small effect. This is problematic when it comes to research on elite athletes, who are, by definition, rare. This problem was encountered in a GWAS carried out in multiple cohorts of elite endurance athletes by Rankinen and colleagues [37]. The authors utilised a cohort of elite Caucasian endurance runners ( $n = 315$ ) and controls, along with a cohort of elite Japanese runners ( $n = 60$ ) and controls, for the discovery phase of the GWAS. Following this discovery phase, in which no SNP met genome-wide significance, forty-two suggestive SNPs were taken through to a replication phase involving endurance athletes and controls from seven other countries, and from different ethnicities. Again, no genome-wide significant SNPs were found, although upon meta-analyses, a single SNP, rs558129 in polypeptide N-Acetylgalactosaminyltransferase 6 (*GALNTL6*) gene, was significant, with the T allele less frequent in athletes than in controls. As such, the authors summarised that there appeared to be no common SNP associated with elite endurance athlete status across this cohort, although they acknowledged their low sample size as a major limiting factor. This limitation is difficult to overcome and represents a significant roadblock in the search for genetic variants associated with elite performance. One potential method is to conduct a GWAS-approach analysis in a cohort of interest, and then attempt to replicate the suggestive findings (i.e., those of suggestive, but not genome-wide, statistical significance) of this analysis in additional, hold-out cohorts. Whilst novel, this approach has only been used in few sport genetics studies [38,39], despite its potential promise. A further potential roadblock is that there may be different associations between SNPs and elite performance across ethnicities, such that a SNP may be performance enhancing in Caucasians, but not in East Asians, for example, thereby requiring the development of ethnicity-specific SNP panels for the purpose of talent identification.

### 3.2. Rare Variants

A further avenue for exploration is that of rare gene variants that may predispose to elite performance. Generally, research focuses on common variants present in  $>1\%$  of the population (i.e., polymorphisms) [2]. However, there are a few identified genetic variants which are very rare, and yet have the potential to positively affect performance. One such variant occurs at rs121917830 within the erythropoietin receptor (*EPOR*) gene. This variation is linked to a disease called erythrocytosis-1, where affected people have increased red blood cell production, and hence greater oxygen carrying capacity [2]. This outcome is potentially advantageous for endurance sports, and in fact one elite athlete, the Finnish cross-country skiing champion Eero Mäntyranta, is known to possess this variant [40]. Preliminary data suggest an additional example of a rare variant, this time affecting the lamin A/C (*LMNA*) gene—related to muscular dystrophy—that was found in Canadian sprint hurdler Priscilla Lopes-Schliep [41]. Finally, a rare variation in the myostatin gene (*MSTN*) has been reported, in which the carrier was described as “extraordinarily muscular” [42]. Such a variation would perhaps be advantageous in sports/events demanding high levels of strength and muscle mass. One issue with the exploration of rare performance-enhancing variants is that, given their very low frequency, they can be hard to identify. In addition, studies of rare variants would need very large samples of unrelated individuals to correctly identify the influence of the variant on athletic performance. Lastly, some variants might also predispose to disease states; a factor raising complex

moral and ethical questions. However, despite these potential issues, research continues towards their identification [41], not least because identification of individuals with disease-causing variants but presenting with a good health status could provide information relating to the underpinning mechanisms of these diseases and provide insight into remedial and resilience-building strategies [43].

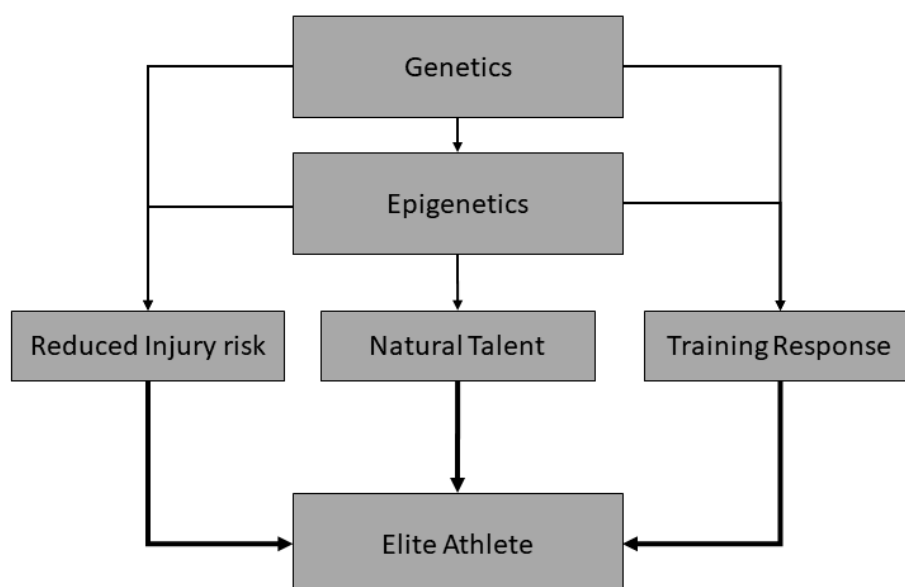
### 3.3. A Predictive Threshold

An area of potential previous confusion is the number of performance-enhancing polymorphisms an individual may require before they are capable, from a genetic standpoint, of elite performance. Both Williams and Folland [34] and Hughes and colleagues [35] report that, in panels containing 23 (endurance) and 22 (strength–power) polymorphisms, the chances of one athlete possessing all was vanishingly small. As such, it seems unlikely that a single athlete possesses the “perfect” genetic profile, and it could be argued that having such a perfect profile is arguably unnecessary. Instead, athletes will likely possess a given number of performance-enhancing polymorphisms. Crucially, the polymorphisms possessed will differ between athletes, such that there might be only minimal overlap between individuals. In this way, once a large number of SNPs responsible for driving elite athlete status are uncovered—should such discovery ever occur—there will be the potential for the development of a threshold number, whereby possession of polymorphisms above this number would be associated with elite performance. Accordingly, there will not necessarily be a commonality in terms of the genetic variants present, although some crossover will certainly occur; instead, the main driving factor will be the total frequency of performance-driving variants. These SNPs will also likely differ between ethnicities, and so ethnicity-specific thresholds and genetic panels will be required. The utilisation of a large number of SNPs reduces the reliance on individual SNPs that occur either at high frequencies across populations (such as the *ACTN3* R allele), or those that, whilst linked to elite athlete status, are still present at relatively low frequencies in elite cohorts (for example the *NRF2* G allele).

### 3.4. Epigenetic Modifications

In addition to genetic variations that might predispose, at least partially, to elite performance in a given sports, favourable epigenetic modifications that predispose to a natural talent trait (direct effect), and/or enhanced response to physical training and reduced risk of injury (indirect effect; (Figure 2)) would also likely be advantageous. In this context, epigenetic modifications—changes in genetic expression that are not due to variations in the underlying genetic code—may modify the attainment of elite athlete status [44] and represent another source of heritable variation that might help in sports talent identification [2]. These epigenetic changes are typically comprised of DNA methylation, histone modifications, and non-coding RNAs, particularly microRNAs (miRNAs) [44,45]. For example, miRNAs are involved in post-transcriptional control of messenger RNA (mRNA) and they include a highly complex network that regulates protein expression within the cell. Sequence and functional conservation of many human miRNAs with distantly related organisms suggests their crucial role in cellular progress [46]. Given their fundamental role in cell functioning, miRNAs have been shown to influence both skeletal muscle differentiation [47] and the magnitude of exercise-induced adaptations [48], which are important components of the journey towards elite athlete status. The use of epigenetic information might help to discover individuals with high chances of possessing characteristics associated with elite sport performance, but also to understand what and how epigenetic modifications can be modulated through appropriate training and dietary regimes to improve performance [2]. As of yet, whilst it appears that epigenetic changes may modify the response to exercise [44,45], it is not clear what overall effect they may have on the attainment of elite status. Additionally, such modifications have the potential to be passed down through generations [49], and thus might form part of the heritable aspect of elite athlete status, although the evidence that this actually occurs in humans is currently weak. Nevertheless, the ability to test for these changes, which are often tissue-specific, could assist in the identification of talented athletes; however, the quantification of epigenetic modifications can, in some cases, be challenging. Many

epigenetic modifications are often both tissue specific and transient [50]; as a result, quantification would require frequent tissue sampling, which in the case of muscle would be invasive and traumatic, and in the case of other tissue, such as the heart and brain, highly impractical [51]. However, early research suggests that salivary profiling of DNA methylation markers holds promise [52], making this approach potentially more attractive and practically realistic.



**Figure 2.** Influence of genetics and epigenetics on traits associated with elite performance. To succeed in sport, an athlete must possess genetic and epigenetic variations that might predispose to a natural talent trait (direct effect), and/or to enhanced response to physical training, and/or to reduced risk of injury (indirect effect).

### 3.5. Lessons from Disease Prediction

One area where the use of genetics to make informed predictions of a future event has been well explored is that of disease risk. Whilst some diseases, such as cystic fibrosis, are monogenic, most diseases are complex and polygenic in nature [53]. Similar to elite athlete status, whilst many diseases have been shown to have a large genetic component, the disease-causing variants identified to date explain little of the variance between individuals; an issue referred to as the ‘missing heritability’ problem [54]. One suggested approach to overcome the problem of missing heritability is to lower the threshold for discovery of SNPs affecting the trait of interest. Due to the high number of comparisons carried out in a GWAS, statistical significance for discovery of new variants is set at  $p < 5 \times 10^{-8}$ . However, the lowering of this threshold has been shown to lead to explanation of a greater proportion of heritability [55]. Recently, Khera and colleagues [56] utilised a TGS comprised of 6,630,150 polymorphisms to create a risk score for coronary artery disease that had an area under the curve of 0.81, suggesting a strong predictive ability; many of these polymorphisms had miniscule effect sizes and weak significance, and yet combined to produce a powerful predictive tool. Such a method clearly holds promise for traits that have a large—but poorly elucidated—genetic component [57], such as elite athlete status. Indeed, returning to the recent GWAS on elite endurance status, whilst no SNP was discovered at the genome-wide significance level ( $p < 5 \times 10^{-8}$ ), a number of SNPs had suggestive significance, and so may hold predictive ability as part of a TGS. As such, it appears that, in order to successfully predict future elite athlete status, models involving multiple genetic variants with low effect sizes are likely required, although this approach has recently been criticised from a disease-prediction viewpoint [58]. However, we return to the common issue of sample size; for discovery of relatively common genetic variants with small effect sizes, sample sizes in excess of



10,000 subjects are likely required [59]—a number considerably greater than that of all truly elite athletes within a specific performance trait on the planet.

#### 4. Even If We Could Test for Talent—Should We?

There are serious and well-placed concerns about the use of genetic information, either alone or in combination with existing measures, for talent identification within sport [21]. It is generally considered that, within sporting contexts, genetic testing should not be carried out on athletes younger than 18 years of age [21]. If genetic testing for talent does become evidence-based, then there will be a requirement for the development of guidelines on its use, in part to protect vulnerable young children. For example, should sports clubs be able to demand players undergo genetic testing as part of their talent identification programmes? Who should have access to the genetic data? Will it be used in a discriminatory way? What if a player refuses to undergo a genetic test? What happens if testing uncovers a potential disease-causing variant? The latter point is potentially an important issue, as whilst it could lead to health-promoting medical interventions, it could also lead to unnecessary medical explorations, as well as increased worry, both on the part of the athlete tested and their relatives, who may also carry the disease-causing variant. Such outcomes are highly unpalatable, and likely represent an extreme example, but demonstrate the potential misuse of genetic information for talent identification.

If and when genetic testing is used to predict future elite athlete status, there will be many false positives and false negatives (i.e., many individuals will be misattributed to future elite or non-elite status [60]). Whilst such error rates may be acceptable at a population level, they are obviously ethically troubling at an individual one—who has the right to tell a young athlete that they do not have the genetic potential to succeed in a given sport? Perhaps more importantly, what impact would receiving this information have on that individual's future exercise behaviour, which, given the wide-ranging health benefits of exercise [61], is an important consideration for lifelong health. For these reasons, some authors have suggested that emphasis in the use of genetic testing for talent identification in children needs to be focused on the wellbeing on the child in the present, not on the potential of future benefits that may never arrive [22].

#### 5. What Could We Potentially Use Genetic Testing for?

It has previously been argued that, instead of using genetic information for talent identification in the traditional sense, this information could be used to match individuals to the type of training to which they are most suited, and from which they will elicit the greatest adaptations [62–64]. Additionally, genetic information could be utilised to identify those athletes with an increased risk of injury, allowing for the provision of pre-emptive strategies to reduce that risk. For example, Varley and colleagues [65] identified a number of polymorphisms associated with an increased stress fracture risk in a cohort of elite athletes; in this case, the high-risk athletes could undergo additional bone mineral density monitoring, along with targeted interventions, such as vitamin D and calcium supplementation.

Whilst traditional research has focused on the physiological drivers of elite athlete status (reviewed in detail by Ahmetov and colleagues [16]), there is the potential to explore the genetic underpinnings of psychological factors, such as anxiety, stress resilience, and skill acquisition. This has not yet been explored in detail, although a SNP in catechol-O-methyltransferase (*COMT*) gene, rs4680, has been linked with competition performance in swimmers [66] and with personality traits in ultra-endurance athletes [67]. Catechol-o-methyltransferase plays a role in the regulation of dopamine within the prefrontal cortex [19]; variation in this SNP affects dopamine levels and thus information processing and memory [19]. Emerging research has also shown an association between a number of polymorphisms and the skill acquisition process [68]. Finally, a number of genetic variants have been linked to an increased susceptibility to concussion injuries [69]. As a result, whilst this information could be used to bias against those with the perceived “unfavourable” genotypes, it could also be used to personalize the training process, identifying those athletes who need greater attention in these areas. Furthermore, with regard to both injury and concussion, the information

could be used to better inform preventative methods, along with increasing the personalisation of recovery and return-to-play protocols, particularly given the evidence that genetic information might enhance adherence [70]. That said, similar to the use of genetic information as a talent identification tool, the research in this area is currently under-developed, and plagued with similar issues to those outlined in this manuscript. The early and initial studies exploring the use of genetic information to training programme design and management, whilst promising, require replication in different cohorts and larger samples to confirm their findings.

## 6. Conclusions

Although there is a well-established effect of genetics on the attainment of elite athlete status, based on the available evidence, it is clear that the current use of genetic tests for the prediction of future elite athlete status is ineffectual, a finding that echoes recent consensus statements [21,71]. Despite the increasing availability of commercial DTC genetic tests, the currently available data suggest the use of the information provided by these tests for talent identification or selection, especially in children, is unfounded. In order to be able to use genetic information within the talent identification process, a far greater number of performance-enhancing polymorphisms needs to be both discovered and replicated in subsequent studies. Future studies in elite athletes must contain detailed information on phenotypes associated with sporting performance, in conjunction with obtaining samples such as blood or muscle tissue that help identify the association between genetics and epigenetics. The combination of performance-enhancing polymorphisms into a TGS, especially if the evidence threshold is lowered, appears to offer a solution to the currently limited predictive capabilities of small numbers of genetic variants. As the evidence base grows, we believe it will be possible to determine a TGS threshold, above which an individual's chance of achieving elite athlete status in a given sport or event is higher. However, we are also of the opinion that there will always be individuals with a score below this threshold who go on to achieve elite athlete status, and those with scores above the threshold who will not be elite athletes. Because elite athlete status is a manifestation of a number of variables, not just genotype, it seems unlikely that it will ever be possible to use genetic information to unequivocally identify a future elite athlete. At best, genetic information may represent a potentially useful adjunct to existing talent identification procedures, enhancing the process of selection, particularly as genetic information is not subject to some of the issues that commonly plague traditional talent identification processes, such as maturation and the relative age effect. Additionally, it has previously been argued that genetic information may be used in the future to identify those with the greatest potential to show favourable adaptations to training [62,64], as well as determine the optimal training type to elicit such adaptations [63]. Furthermore, there is the potential to utilise such information to reduce injury occurrence [72]. Again, and we must be clear; such information should not be used as a standalone, but as an adjunct to current talent identification processes, thereby allowing the training process to become more personalised, and enabling athletes to get ever closer to their maximum potential.

**Author Contributions:** Conceptualisation, C.P. and J.K.; writing—original draft preparation, C.P. and J.K.; writing—review and editing, J.G, J.D.C. and A.L.; supervision, J.K.

**Funding:** This research received no external funding.

**Acknowledgments:** The authors of this review want to acknowledge all the investigators that have contributed to determine the association of genetics and athletic performance.

**Conflicts of Interest:** Craig Pickering is a former employee of DNAFit Life Sciences, a genetic testing company. He received no financial incentives for the preparation of this manuscript, which was produced as part of his Doctoral studies. John Kiely, Jozo Grgic, Juan Del Coso and Alejandro Lucia declare no conflict of interest relevant to the content of this article.

## References

1. Williams, A.M.; Reilly, T. Talent identification and development in soccer. *J. Sports Sci.* **2000**, *18*, 657–667.

2. Moran, C.N.; Pitsiladis, Y.P. Tour de France Champions born or made: Where do we take the genetics of performance? *J. Sports Sci.* **2017**, *35*, 1411–1419.
3. De Moor, M.H.; Spector, T.D.; Cherkas, L.F.; Falchi, M.; Hottenga, J.J.; Boomsma, D.I.; De Geus, E.J. Genome-wide linkage scan for athlete status in 700 British female DZ twin pairs. *Twin Res. Hum. Genet.* **2007**, *10*, 812–820.
4. Houweling, P.J.; Papadimitriou, I.D.; Seto, J.T.; Pérez, L.M.; Coso, J.D.; North, K.N.; Lucia, A.; Eynon, N. Is evolutionary loss our gain? The role of ACTN3 p. Arg577Ter (R577X) genotype in athletic performance, ageing, and disease. *Hum. Mutat.* **2018**, *39*, 1774–1787.
5. Ma, F.; Yang, Y.; Li, X.; Zhou, F.; Gao, C.; Li, M.; Gao, L. The association of sport performance with ACE and ACTN3 genetic polymorphisms: A systematic review and meta-analysis. *PLoS ONE* **2013**, *8*, e54685.
6. Vincent, B.; De Bock, K.; Ramaekers, M.; Van den Eede, E.; Van Leemputte, M.; Hespel, P.; Thomis, M.A. ACTN3 (R577X) genotype is associated with fiber type distribution. *Physiol. Genomics* **2007**, *32*, 58–63.
7. Yang, N.; MacArthur, D.G.; Gulbin, J.P.; Hahn, A.G.; Beggs, A.H.; Eastale, S.; North, K. ACTN3 genotype is associated with human elite athletic performance. *Am. J. Hum. Genet.* **2003**, *73*, 627–631.
8. Ahmetov, I.I.; Druzhevskaya, A.M.; Lyubaeva, E.V.; Popov, D.V.; Vinogradova, O.L.; Williams, A.G. The dependence of preferred competitive racing distance on muscle fibre type composition and ACTN3 genotype in speed skaters. *Exp. Physiol.* **2011**, *96*, 1302–1310.
9. Eynon, N.; Sagiv, M.; Meckel, Y.; Duarte, J.A.; Alves, A.J.; Yamin, C.; Sagiv, M.; Goldhammer, E.; Oliveira, J. NRF2 intron 3 A/G polymorphism is associated with endurance athletes' status. *J. Appl. Physiol.* **2009**, *107*, 76–79.
10. Scott, R.A.; Irving, R.; Irwin, L.; Morrison, E.; Charlton, V.; Austin, K.; Tladi, D.; Deason, M.; Headley, S.A.; Kolkhorst, F.W.; et al. ACTN3 and ACE genotypes in elite Jamaican and US sprinters. *Med. Sci. Sports Exerc.* **2010**, *42*, 107–112.
11. Moran, C.N.; Yang, N.; Bailey, M.E.; Tsiokanos, A.; Jamurtas, A.; MacArthur, D.G.; North, K.; Pitsiladis, Y.P.; Wilson, R.H. Association analysis of the ACTN3 R577X polymorphism and complex quantitative body composition and performance phenotypes in adolescent Greeks. *Eur. J. Hum. Genet.* **2007**, *15*, 88.
12. Papadimitriou, I.D.; Lucia, A.; Pitsiladis, Y.P.; Pushkarev, V.P.; Dyatlov, D.A.; Orekhov, E.F.; Artioli, G.G.; Guilherme, J.P.; Lancha, A.H.; Ginevičienė, V.; et al. ACTN3 R577X and ACE I/D gene variants influence performance in elite sprinters: A multi-cohort study. *BMC Genomics* **2016**, *17*, 285.
13. Berman, Y.; North, K.N. A gene for speed: The emerging role of  $\alpha$ -actinin-3 in muscle metabolism. *Physiology* **2010**, *25*, 250–259.
14. MacArthur, D.G.; North, K.N. A gene for speed? The evolution and function of  $\alpha$ -actinin-3. *Bioessays* **2004**, *26*, 786–795.
15. Lee, F.X.; Houweling, P.J.; North, K.N.; Quinlan, K.G. How does  $\alpha$ -actinin-3 deficiency alter muscle function? Mechanistic insights into ACTN3, the 'gene for speed'. *Biochim. Biophys. Acta* **2016**, *1863*, 686–693.
16. Ahmetov, I.I.; Egorova, E.S.; Gabdrakhmanova, L.J.; Fedotovskaya, O.N. Genes and athletic performance: An update. *Med. Sports Sci.* **2016**, *61*, 41–54.
17. Silventoinen, K.; Sammalisto, S.; Perola, M.; Boomsma, D.I.; Cornes, B.K.; Davis, C.; Dunkel, L.; De Lange, M.; Harris, J.R.; Hjelmborg, J.V.; et al. Heritability of adult body height: A comparative study of twin cohorts in eight countries. *Twin Res.* **2003**, *6*, 399–408.
18. Elks, C.E.; Den Hoed, M.; Zhao, J.H.; Sharp, S.J.; Wareham, N.J.; Loos, R.J.; Ong, K.K. Variability in the heritability of body mass index: A systematic review and meta-regression. *Front. Endocrinol.* **2012**, *3*, 29.
19. Stein, M.B.; Fallin, M.D.; Schork, N.J.; Gelernter, J. COMT polymorphisms and anxiety-related personality traits. *Neuropsychopharmacology* **2005**, *30*, 2092.
20. Peplonska, B.; Safranow, K.; Adamczyk, J.; Boguszewski, D.; Szymański, K.; Soltyszewski, I.; Barczak, A.; Siewierski, M.; Ploski, R.; Sozanski, H.; et al. Association of serotonergic pathway gene variants with elite athletic status in the Polish population. *J. Sports Sci.* **2019**, doi:10.1080/02640414.2019.1583156.
21. Webbarn, N.; Williams, A.; McNamee, M.; Bouchard, C.; Pitsiladis, Y.; Ahmetov, I.; Ashley, E.; Byrne, N.; Camporesi, S.; Collins, M.; et al. Direct-to-consumer genetic testing for predicting sports performance and talent identification: Consensus statement. *Br. J. Sports Med.* **2015**, *49*, 1486–1491.
22. Camporesi, S.; McNamee, M.J. Ethics, genetic testing, and athletic talent: children's best interests, and the right to an open (athletic) future. *Physiol. Genomics* **2016**, *48*, 191–195.
23. Williams, A.G.; Heffernan, S.M.; Day, S.H. Genetic testing in exercise and sport-have direct-to-consumer genetic tests come of age? *Sci. Sport Curr. Trends* **2014**, *2*, 3–9.

24. Gayagay, G.; Yu, B.; Hambly, B.; Boston, T.; Hahn, A.; Celermajer, D.S.; Trent, R.J. Elite endurance athletes and the ACE I allele—The role of genes in athletic performance. *Hum. Genet.* **1998**, *103*, 48–50.
25. Lucia, A.; Gómez-Gallego, F.; Barroso, I.; Rabadán, M.; Bandrés, F.; San Juan, A.F.; Chicharro, J.L.; Ekelund, U.; Brage, S.; Earnest, C.P.; et al. PPARGC1 genotype (Gly482Ser) predicts exceptional endurance capacity in European males. *J. Appl. Physiol.* **2005**, *99*, 344–348.
26. Maciejewska, A.; Sawczuk, M.; Cieszczyk, P.; Mozhayskaya, I.A.; Ahmetov, I.I. The PPARGC1A gene Gly482Ser in Polish and Russian athletes. *J. Sports Sci.* **2012**, *30*, 101–113.
27. Ahmetov, I.I.; Popov, D.V.; Missina, S.S.; Vinogradova, O.L.; Rogozkin, V.A. Association of mitochondrial transcription factor (TFAM) gene polymorphism with physical performance in athletes. *Hum. Physiol.* **2010**, *36*, 229–233.
28. Williams, A.G.; Wackerhage, H.; Day, S.H. Genetic testing for sports performance, responses to training and injury risk: Practical and ethical considerations. *Med. Sport Sci.* **2016**, *61*, 105–119.
29. Yang, N.; Macarthur, D.G.; Wolde, B.; Onywera, V.O.; Boit, M.; Lau, S.Y.; Wilson, R.H.; Scott, R.A.; Pitsiladis, Y.; North, K.N. The ACTN3 R577X polymorphism in East and West African athletes. *Med. Sci. Sports Exerc.* **2007**, *39*, 1985–1988.
30. Lucia, A.; Oliván, J.; Gómez-Gallego, F.; Santiago, C.; Montil, M.; Foster, C. Citius and longius (faster and longer) with no  $\alpha$ -actinin-3 in skeletal muscles? *Br. J. Sports Med.* **2007**, *41*, 616–617.
31. Ruiz, J.R.; Gómez-Gallego, F.; Santiago, C.; González-Freire, M.; Verde, Z.; Foster, C.; Lucia, A. Is there an optimum endurance polygenic profile? *J. Physiol.* **2009**, *587*, 1527–1534.
32. Ruiz, J.R.; Arteta, D.; Buxens, A.; Artieda, M.; Gómez-Gallego, F.; Santiago, C.; Yvert, T.; Morán, M.; Lucia, A. Can we identify a power-oriented polygenic profile? *J. Appl. Physiol.* **2010**, *108*, 561–566.
33. Santiago, C.; Ruiz, J.R.; Muniesa, C.A.; González-Freire, M.; Gómez-Gallego, F.; Lucia, A. Does the polygenic profile determine the potential for becoming a world-class athlete? Insights from the sport of rowing. *Scand. J. Med. Sci. Sports* **2010**, *20*, e188–e194.
34. Williams, A.G.; Folland, J.P. Similarity of polygenic profiles limits the potential for elite human physical performance. *J. Physiol.* **2008**, *586*, 113–121.
35. Hughes, D.C.; Day, S.H.; Ahmetov, I.I.; Williams, A.G. Genetics of muscle strength and power: Polygenic profile similarity limits skeletal muscle performance. *J. Sports Sci.* **2011**, *29*, 1425–1434.
36. Rodas, G.; Osaba, L.; Arteta, D.; Pruna, R.; Fernández, D.; Lucia, A. Genomic prediction of tendinopathy risk in elite team sports. *Int. J. Sports Physiol. Perform.* **2019**, doi:10.1123/ijspp.2019-0431.
37. Rankinen, T.; Fuku, N.; Wolfarth, B.; Wang, G.; Sarzynski, M.A.; Alexeev, D.G.; Ahmetov, I.I.; Boulay, M.R.; Cieszczyk, P.; Eynon, N.; et al. No evidence of a common DNA variant profile specific to world class endurance athletes. *PLoS ONE* **2016**, *11*, e0147330.
38. Ahmetov, I.I.; Kulemin, N.A.; Popov, D.V.; Naumov, V.A.; Akimov, E.B.; Bravy, Y.R.; Egorova, E.S.; Galeeva, A.A.; Generozov, E.V.; Kostryukova, E.S.; et al. Genome-wide association study identifies three novel genetic markers associated with elite endurance performance. *Biol. Sport* **2015**, *32*, 3–9.
39. Pickering, C.; Suraci, B.; Semenova, E.A.; Boulygina, E.A.; Kostryukova, E.S.; Kulemin, N.A.; Borisov, O.V.; Khabibova, S.A.; Larin, A.K.; Pavlenko, A.V.; et al. A Genome-Wide Association Study of sprint performance in elite youth football players. *J. Strength Cond. Res.* **2019**, *33*, 2344–2351.
40. Juvonen, E.; Ikkala, E.; Fyhrquist, F.; Ruutu, T. Autosomal dominant erythrocytosis caused by increased sensitivity to erythropoietin. *Blood* **1991**, *78*, 3066–3069.
41. Waggott, D.; Mattsson, C.M.; Wheeler, M.; Ashley, E.A. The Genomics of Extreme Athletes. The ELITE Study (Exercise at the Limit-Inherited Traits of Endurance) [Abstract]. In Proceedings of the American Society of Human Genetics (ASHG) Annual Meeting, Vancouver, BC, Canada, 18–22 October 2016.
42. Schuelke, M.; Wagner, K.R.; Stolz, L.E.; Hübner, C.; Riebel, T.; Kömen, W.; Braun, T.; Tobin, J.F.; Lee, S.J. Myostatin mutation associated with gross muscle hypertrophy in a child. *N. Engl. J. Med.* **2004**, *350*, 2682–2688.
43. Chen, R.; Shi, L.; Hakenberg, J.; Naughton, B.; Sklar, P.; Zhang, J.; Zhou, H.; Tian, L.; Prakash, O.; Lemire, M.; et al. Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases. *Nat. Biotechnol.* **2016**, *34*, 531.
44. Ehler, T.; Simon, P.; Moser, D.A. Epigenetics in sports. *Sports Med.* **2013**, *43*, 93–110.
45. Voisin, S.; Eynon, N.; Yan, X.; Bishop, D.J. Exercise training and DNA methylation in humans. *Acta Physiol.* **2015**, *213*, 39–59.

46. Ghosh, A.; Mahajan, P.B. Can genotype determine the sports phenotype? A paradigm shift in sports medicine. *J. Basic Clin. Physiol. Pharmacol.* **2016**, *27*, 333–339.
47. Luo, W.; Nie, Q.; Zhang, X. MicroRNAs involved in skeletal muscle differentiation. *J. Genet. Genomics* **2013**, *40*, 107–116.
48. Davidsen, P.K.; Gallagher, I.J.; Hartman, J.W.; Tarnopolsky, M.A.; Dela, F.; Helge, J.W.; Timmons, J.A.; Phillips, S.M. High responders to resistance exercise training demonstrate differential regulation of skeletal muscle microRNA expression. *J. Appl. Physiol.* **2011**, *110*, 309–317.
49. Richards, E.J. Inherited epigenetic variation—Revisiting soft inheritance. *Nat. Rev. Genet.* **2006**, *7*, 395.
50. Lekk, K.; Modhukur, V.; Rajashekar, B.; Märtens, K.; Mägi, R.; Kolde, R.; Koltšina, M.; Nilsson, T.K.; Vilo, J.; Salumets, A.; et al. DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biol.* **2014**, *15*, 324.
51. Pickering, C.; Kiely, J. The Development of a Personalised Training Framework: Implementation of Emerging Technologies for Performance. *J. Funct. Morphol. Kinesiol.* **2019**, *4*, 25.
52. Langie, S.A.; Moisse, M.; Declerck, K.; Koppen, G.; Godderis, L.; Vanden Berghe, W.; Drury, S.; De Boever, P. Salivary DNA methylation profiling: Aspects to consider for biomarker identification. *Basic Clin. Pharmacol. Toxicol.* **2017**, *121*, 93–101.
53. Chatterjee, N.; Shi, J.; García-Closas, M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat. Rev. Genet.* **2016**, *17*, 392.
54. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorf, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A.; et al. Finding the missing heritability of complex diseases. *Nature* **2009**, *461*, 747.
55. Shi, H.; Kichaev, G.; Pasaniuc, B. Contrasting the genetic architecture of 30 complex traits from summary association data. *Am. J. Hum. Genet.* **2016**, *99*, 139–153.
56. Khera, A.V.; Chaffin, M.; Aragam, K.G.; Haas, M.E.; Roselli, C.; Choi, S.H.; Natarajan, P.; Lander, E.S.; Lubitz, S.A.; Ellinor, P.T.; et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **2018**, *50*, 1219.
57. Dudbridge, F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet.* **2013**, *9*, e1003348.
58. Janssens, A.C.; Joyner, M.J. Polygenic risk scores that predict common diseases using millions of single nucleotide polymorphisms: Is more, better? *Clin. Chem.* **2019**, doi:10.1373/clinchem.2018.296103.
59. Mattsson, C.M.; Wheeler, M.T.; Waggott, D.; Caleshu, C.; Ashley, E.A. Sports genetics moving forward: Lessons learned from medical research. *Physiol. Genomics* **2016**, *48*, 175–182.
60. Breitbach, S.; Tug, S.; Simon, P. Conventional and genetic talent identification in sports: Will recent developments trace talent? *Sports Med.* **2014**, *44*, 1489–1503.
61. Pareja-Galeano, H.; Garatachea, N.; Lucia, A. Exercise as a polypill for chronic diseases. *Prog. Mol. Biol. Transl. Sci.* **2015**, *135*, 497–526.
62. Pickering, C.; Kiely, J. Can the ability to adapt to exercise be considered a talent—And if so, can we test for it? *Sports Med. Open* **2017**, *3*, 43.
63. Jones, N.; Kiely, J.; Suraci, B.; Collins, D.J.; De Lorenzo, D.; Pickering, C.; Grimaldi, K.A. A genetic-based algorithm for personalized resistance training. *Biol. Sport* **2016**, *33*, 117.
64. Pickering, C.; Kiely, J.; Suraci, B.; Collins, D. The magnitude of Yo-Yo test improvements following an aerobic training intervention are associated with total genotype score. *PLoS ONE* **2018**, *13*, e0207597.
65. Varley, I.; Hughes, D.C.; Greeves, J.P.; Stellingwerff, T.; Ranson, C.; Fraser, W.D.; Sale, C. The association of novel polymorphisms with stress fracture injury in elite athletes: Further insights from the SFEA cohort. *J. Sci. Med. Sport* **2018**, *21*, 564–568.
66. Abe, D.; Doi, H.; Asai, T.; Kimura, M.; Wada, T.; Takahashi, Y.; Matsumoto, T.; Shinohara, K. Association between COMT Val158Met polymorphism and competition results of competitive swimmers. *J. Sports Sci.* **2018**, *36*, 393–397.
67. van Breda, K.; Collins, M.; Stein, D.J.; Rauch, L. The COMT val158met polymorphism in ultra-endurance athletes. *Physiol. Behav.* **2015**, *151*, 279–283.
68. Jacob, Y.; Cripps, A.; Evans, T.; Chivers, P.T.; Joyce, C.; Anderton, R.S. Identification of genetic markers for skill and athleticism in sub-elite Australian football players: A pilot study. *J. Sports Med. Phys. Fit.* **2018**, *58*, 241–248.

69. McFie, S.; Abrahams, S.; Patricios, J.; Suter, J.; Posthumus, M.; September, A.V. The association between COMT rs4680 and 5-HTTLPR genotypes and concussion history in South African rugby union players. *J. Sports Sci.* **2018**, *36*, 920–933.
70. Nielsen, D.E.; El-Sohemy, A. Disclosure of genetic information and change in dietary intake: A randomized controlled trial. *PLoS ONE* **2014**, *9*, e112665.
71. Vlahovich, N.; Fricker, P.A.; Brown, M.A.; Hughes, D. Ethics of genetic testing and research in sport: A position statement from the Australian Institute of Sport. *Br. J. Sports Med.* **2017**, *51*, 5–11.
72. Heffernan, S.M.; Kilduff, L.P.; Day, S.H.; Pitsiladis, Y.P.; Williams, A.G. Genomics in rugby union: A review and future prospects. *Eur. J. Sport Sci.* **2015**, *15*, 460–468.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).